

The Importance of Teachers

Ben Backes

K-12 Education

Last updated: December 2025

Introduction

More than 4 million individuals work as schoolteachers in the United States.¹ Collectively, teachers in public schools represent more than 2% of the American workforce² and command over 1% of U.S. GDP in salary.³ As the primary nonfamily adult contacts for school-aged children, teachers play an important role in shaping the next generation and are arguably the primary way through which schools impact students.⁴

Public investment in teachers is not limited to financial outlays: state and district education agencies invest time crafting myriad policies that influence who is permitted to become and remain a teacher. For instance, at the state level, key policy levers affecting prospective teachers include requirements for teacher preparation programs where prospective teachers receive their training, requirements for student teaching, and requirements for receiving an initial teaching license, such as passing general content and subject matter tests. States also dictate guidelines for teacher performance evaluations and mandate professional development hour requirements for renewing teaching licenses.⁵ At the district level, there tends to be considerable freedom in how teaching evaluations are conducted within the frameworks established by states.⁶

The introduction of annual testing in grades 3–8 mandated by No Child Left Behind, paired with the development of state and district databases that allow students to be linked to their teachers and tracked over time, spurred an explosion of research on the impacts of teachers on their students. Much of this research focused on the effects of teachers on student test growth, but a newer literature also assesses the impact of teachers on nontest K-12 schooling outcomes, such as attendance, course-taking, and high school graduation. Linkages to external databases outside typical state and district data collection have also allowed researchers to further explore the impact of teachers on student outcomes that extend beyond K-12 schooling, such as post-secondary enrollment, interactions with the criminal justice system, and adult labor market participation and earnings. Unsurprisingly, given the amount of time that teachers spend with their students, the empirical evidence has demonstrated teacher impacts on all these and other short- and long-run student outcomes.

Key findings

Key finding #1: *There are educationally meaningful differences between teachers in terms of their ability to help their students learn, as measured by tests.*

After accounting for students' starting points, as measured by prior test scores, some teachers consistently raise students' test scores in math and English language arts (ELA) significantly more than others. In math, the difference between a teacher at the 25th and 75th percentiles is similar to the expected improvement corresponding to a 10-student reduction in class size. Differences across teachers are estimated to be larger in math than in ELA and larger in elementary school than in secondary school.

Key finding #2: *Teachers also impact nontest student outcomes that can be measured with typical state and district data collection, such as school days missed due to absences and suspensions.*

This across-teacher variation in improving the nontest outcomes of students—used as a proxy for underlying student soft skills such as self-discipline—has been found in outcomes such as days missed due to absences, days missed due to disciplinary incidents, grade progression, and (in secondary school) course grades. The across-teacher variation in impacts on nontest outcomes is similar to the across-teacher variation in impacts on test outcomes.

Key finding #3: *There is a weak relationship between teacher impacts on test scores and teacher impacts on nontest outcomes.*

This finding suggests that test score impacts and nontest impacts capture different underlying facets of teacher skill. Hence, a narrow focus on test scores misses other important and measurable ways in which teachers impact their students.

Key finding #4: *Better teachers improve their students' long-run student outcomes.*

These long-run outcomes include high school graduation, college-going, and earnings in adulthood. While the evidence base is stronger for impacts on test scores translating into improvements in long-run outcomes, an emerging literature has also found long-run effects operating through nontest outcomes, especially for students on the margin of graduating from high school or attending college.

Key finding #5: *Readily quantifiable teacher characteristics capture very little of the variation in teachers' impacts on their students.*

For example, as measured by impacts on student test scores, only 3% of the variation in teacher effectiveness is captured by readily observable teacher measures such as advanced degree attainment and experience.

Understudied topics.

We know little about what teachers are actually doing inside or outside the classroom that leads to differences in their impacts on students. Which teacher practices are

supportive of their students learning more in math or of being less likely to be chronically absent from school? Additionally, is it possible to design interventions that consistently improve these practices or to know before the hiring stage which teachers are stronger in these skills?

Policy considerations.

Teacher impacts on test scores are one of the most frequently studied topics in education research. As such, most of the findings listed above are obtained from multiple studies covering different contexts that have produced similar results. Studies on teacher impacts on nontest outcomes, such as student attendance, involve smaller samples and fewer sites. So, how well the findings generalize across settings is less clear. Finally, there is a disconnect between the evidence on teacher impacts on test and nontest outcomes—where research has found wide variation across teachers—and the evidence on policy-shaped outcomes such as teacher evaluation ratings and pay—where most teachers receive the same rating and where teachers with the same level of experience and the same degree tend to receive the same pay. That is, teacher policies do not consistently measure or respond to the variation in teacher effectiveness.

Evidence

As described in further detail below, estimates of teacher quality are obtained by comparing the outcomes of students who have specific teachers to the outcomes of students with other teachers whose students have similar prior achievement and demographic characteristics and by thus obtaining estimates of variation in effectiveness across teachers. This evidence has been supplemented by the results of a set of randomization studies designed to measure bias in estimates of teacher effectiveness.

The extent to which there is across-teacher variation in teachers' impacts on student outcomes is of policy interest because it suggests whether there is scope for improvement. For instance, if there are significant differences across teachers, then we might wonder which teacher characteristics (e.g., experience) predict this variation and consider policies designed to improve the overall quality of the teacher workforce. Policies might be focused on improving the capacities of individual teachers (e.g., through improved teacher education or professional development), incentivizing higher levels of performance, or changing who is in the teacher workforce.

Key finding #1: *There are educationally meaningful differences between teachers in terms of their ability to help their students learn, as measured by tests.*

As described in more detail in this review,⁷ estimates of teacher impacts on student learning are typically obtained through a regression of test scores in a given year on prior test scores and the observable characteristics of their students, such as free- or reduced-price lunch eligibility, English language learner status, disability status, and gender. This method generates a measure of student learning over what would otherwise be expected, often called “value added” (VA). There is now a broad consensus that VA is a useful tool for capturing how teachers impact their students, and the use of



test scores as a teacher measure is now widespread in both research and policy. However, there is far less consensus about how VA should be used in policy.⁸

The primary concern with VA estimation is the possibility that these estimates are biased due to the process of assigning students to classrooms. For example, we might expect that if certain teachers consistently have low-ability students assigned to their classrooms, we would run the risk of conflating a teacher effect with the assignment of students to teachers and what that portends about their test score growth.

As described in more detail above, there are dozens of studies going back decades that estimate variation in teacher quality, but the most rigorous nonexperimental studies take advantage of the recent spread in availability of state- or districtwide longitudinal data that track students over time. These studies use a quasi-experimental design (QED) to examine the VA of teachers who switch schools or grades to obtain estimates of bias associated with the sorting of students to teachers. The premise of the test is that if a high-VA teacher leaves a school, for example, we would expect a drop in aggregate test scores in the departure school and a rise in the arrival school. In the paper where this design was developed,⁹ the authors examine a panel of nearly 20 years of elementary- and middle-school teachers in New York City and find that changes in school achievement can be accurately predicted by changes in the estimated VA of the teacher workforce. This study has since been replicated¹⁰ in multiple settings.¹¹

Additional evidence on bias in VA estimates comes from a set of studies implementing random assignment of students to classrooms. In these papers, teacher impacts are calculated prior to random assignment and are then compared to estimates following assignment. If VA estimates were unbiased, we would expect to be unable to rule out equality of VA from before and after randomization. Across three studies, this is indeed what we observe once the prior achievement of students in the classroom is taken into account: equality of forecast and actual VA cannot be rejected.¹²

There is one additional finding from the original New York study noted above worth highlighting. The authors link K-12 administrative data to tax records to estimate the scope of bias in VA estimates associated with factors that are typically unobserved in state or district longitudinal data, such as parental income. The authors find that the bias associated with omitted parental background variables is very close to zero once prior test scores are accounted for—a finding attributed to prior achievement being sufficient to remove potential biases from the sorting of students to teachers. This finding is consistent with the QED and randomization studies discussed above.

A pair of additional findings regarding heterogeneous impacts and exposure is worth mentioning. The first is that there is evidence that Black students assigned to Black teachers score higher on standardized tests;¹³ these impacts translate into long-run student outcomes.¹⁴ The second is that disadvantaged students are more likely to be taught by less experienced or less effective teachers,¹⁵ although this finding is not universal, especially when we look within—rather than across—school districts.¹⁶

Finally, the empirical finding of across-teacher variation in effectiveness in raising test scores is supported by survey evidence from both principals and students, which also indicates across-teacher variation in bolstering student learning. For example, a paper



from a midsize school district surveys elementary-school principals in the district and asks them to rate teachers on dimensions such as classroom management and the ability to raise student achievement in math and reading.¹⁷ Not only do principals provide ratings that vary across teachers within their schools, but the principals' ratings of teachers' ability to raise student achievement are also correlated with VA, with correlations on the order of 0.30. Principals appear to be especially adept at distinguishing the very best and worst teachers. Consistent with this evidence from principals, a more recent paper uses student surveys from Massachusetts to link survey responses to teacher VA.¹⁸ These surveys reveal within-school, across-teacher variation in subjective assessments of factors such as the instructional environment, and teachers rated higher by their students on these surveys also produce higher gains in student achievement.

Key finding #2: *Teachers also impact nontest student outcomes that can be measured with typical state and district data collection, such as school days missed due to absences and suspensions.*

A more recent literature has taken the basic framework described above and replaced test scores as the student outcome variable with other outcomes contained in typical administrative data collection, such as days missed due to absences, days missed due to suspensions, grade progression, and (in secondary school) course grades. These estimates of teacher impacts on nontest outcomes are often referred to as “nontest value added” (NVA) to differentiate them from test-based VA described above.

One influential study¹⁹ uses ninth-grade students in North Carolina to estimate NVA. The study first uses factor analysis to reduce the four nontest outcomes noted above into one nontest factor and then estimates teacher impacts on the nontest factor. The paper demonstrates that, as with teacher impacts on test scores, some teachers are systematically better at improving nontest outcomes than others and that the across-teacher variation in nontest impacts is similar in magnitude to the variation in test impacts. However, for a given teacher, the relationship between estimated teacher impacts on test scores and nontest outcomes is weak (correlation = 0.15), suggesting multiple components of teacher effectiveness.

The study described above uses ninth-grade students in one state, but the basic pattern has been found in other settings. Using elementary-school students in North Carolina, another study examines teacher impacts on student absences, again finding systematic across-teacher variation in impacts on absences that is only weakly related to the variation in impacts on test scores.²⁰ An additional study that uses a school district in California finds variation across teachers in both middle and high schools in reducing class-level absences of their students.²¹ A recent study finds that the impacts on suspensions are driven in part by differential rates of office referrals.²²

Relative to test VA, there are some additional considerations that must be taken into account when researchers test for the unbiasedness of teacher nontest estimates. NVA estimation is complicated by the fact that outcomes are not subject specific. We can estimate VA to math for math teachers and VA to ELA for ELA teachers, but for outcomes such as days missed due to absences, each of a student's teachers shares a common outcome, which makes isolating the NVA contribution of teachers more



challenging. For instance, consider a set of students who have two teachers, A and B. We might want to know how Teacher A influences her students, but for the students who have both Teacher A and Teacher B, we worry that the influence of Teacher B on outcomes such as attendance could be misattributed to Teacher A. Another potential concern with NVA is that outcomes such as days missed due to suspensions are influenced by school or district policy to a much greater extent than test scores are.

To validate NVA, research typically uses QEDs in the spirit of the one described above for test scores. In particular, in the North Carolina study, the author shows that the main results are unchanged when instrumented NVA in a school-cohort is used rather than the actual NVA. This result is reassuring because it removes potential biases associated with within-school student-teacher sorting. However, the evidence that NVA can be estimated without bias—i.e., that a one unit change in estimated NVA is predictive of a one unit change in student nontest outcomes across a variety of settings—is much more limited than the evidence for test VA.

Unlike test VA, which has been found to have mostly constant impacts throughout the student achievement distribution, impacts operating through nontest VA tend to be most pronounced for disadvantaged students. The California district study noted above finds that positive effects are especially relevant for students with low achievement and/or low attendance. A study of middle- and high-school students in Massachusetts²³ finds similar results for a composite of nontest outcomes. A natural question is whether these results represent true differences or are a reflection of the fact that these students are more likely to be on the margin for higher attendance, suspension, and related nontest measures. Evidence from the Massachusetts study suggests that the latter may be at least a partial explanation, in that the students whose outcomes are most sensitive to teacher quality tend to be the same students who are on the margin of, for example, receiving any suspension.

Key finding #3: *There is a weak relationship between teacher impacts on test scores and teacher impacts on nontest outcomes.*

The evidence underlying this finding primarily comes from the papers discussed in the NVA section above in regard to the discussion of key finding #2. In particular, these papers construct test VA and nontest VA for each teacher and examine the correlation between test VA and nontest VA. Across the studies cited above, the correlation between test VA and nontest VA ranges from 0.06 to 0.21, depending on the location, grade range(s) covered, and choices made during the VA estimation procedure (e.g., whether to include school fixed effects). Because VA is measured with error, these correlations are attenuated toward zero. As a point of comparison, the correlation between math VA and reading VA for elementary-school teachers is approximately 0.60.²⁴

A closely related paper that uses data from the Measures of Effective Teaching (MET) Project estimates teacher impacts on test scores along with underlying student social-emotional measures such as growth mindset and grit.²⁵ Consistent with the studies above, the paper reports positive correlations between test score impacts and student behavioral impacts, but the correlations are on the order of 0.20.



Taken together, these findings suggest that NVA represents a repeatable teacher measure that captures meaningful impacts on students and is largely uncaptured by test VA. However, very little is known about which teacher skills or practices are related to NVA.

Key finding #4: *Better teachers improve their students' long-run student outcomes.*

Before I cover the evidence on the impact of teachers on long-run student outcomes, I want to point out that a number of papers have estimated the impact of assignment to teachers with higher test and nontest VA on intermediate outcomes in high school.²⁶ These papers generally find that assignment to teachers with higher test VA raises outcomes that are most relevant for students at the top of the achievement distribution, such as SAT scores and taking Advanced Placement (AP) classes and passing AP tests. On the other hand, assignment to teachers with higher nontest VA is more predictive of outcome such as whether a student takes the SAT, is arrested, and graduates from high school.

Turning to the impacts of teachers on long-run student outcomes, I observe that there are two types of evidence underlying this finding. The first is based on directly estimating teacher impacts on long-run student outcomes. For example, another paper using New York City²⁷ data links students' teachers in grades 4–8 to their college-going at age 20 and earnings at age 28. Controlling for parental characteristics, the study finds significant across-teacher variation in impacts on these long-run outcomes. In addition, only a small portion of this variation can be explained by teacher impacts on test scores.

The second type of evidence is derived from estimating the impact of assignment to teachers with higher VA on long-run outcomes. Two additional papers link VA to college-going outcomes. Again using data from New York City, the first²⁸ finds that students taught by high-VA teachers are less likely to have children as teenagers and more likely to attend college, attend a better college, and have higher earnings in adulthood. Estimates suggest that a one standard deviation (s.d.) increase in VA in one grade raises the likelihood of attending college at age 20 by 0.82 percentage points, relative to a sample mean of 37%. Second, the Massachusetts paper originally noted above links secondary school data to college-going data to estimate the long-run impacts of both test and nontest VA.²⁹ While the paper finds test VA to be associated with increases in college quality, it also finds nontest VA to be predictive of college-going and four-year college-going. This latter finding also holds when test VA is controlled for, again indicating that test VA and nontest VA largely capture separate teaching skills that each translate into better student outcomes.

Key finding #5: *Readily quantifiable teacher characteristics capture very little of the variation in teachers' impacts on their students.*

An early study found that 3% of the variation across teachers can be explained by readily observable factors such as advanced degree attainment and experience.³⁰ These are the primary factors upon which teacher pay scales are typically based. Another study found that even when incorporating survey data on typically unobserved factors

such as personality traits and feelings of self-efficacy, only 12 percent of the variation in teacher quality can be explained.³¹

This is not to say that observable measures are completely unrelated to teacher effectiveness. One estimate places the average difference between teachers on the opposite extremes of the observables distribution -- a teacher with 10 years of experience, who went to a competitive undergraduate college, had average licensure test scores, who is licensed, and who is National Board Certified versus one with no experience, who attended a non-competitive undergraduate college, whose licensure scores were 1 standard deviation below average, and who is not National Board Certified -- as equivalent to about the impact of being exposed to a teacher 1-2 standard deviations higher in the underlying teacher quality distribution.³² However, even for teachers with similar levels of experience or other observable characteristics, there is tremendous variation in effectiveness.

There is a large³³ literature³⁴ estimating, for example, returns³⁵ to teaching experience, finding that, *on average*, more experienced teachers are more effective. For a full discussion of the returns to teaching experience, readers are encouraged to visit the review linked here.⁴⁰ Papers that use these approaches to estimate within-teacher returns to experience, as measured by impacts on test scores, overwhelmingly find that the majority of improvement occurs in the first 3–5 years on the job. A handful of papers also examines returns to experience for nontest VA. A paper⁴¹ examining elementary-school teachers in North Carolina finds that students of teachers with twenty years of experience have approximately 0.20 s.d. fewer days missed due to absences. Another paper⁴² examining math and ELA middle-school teachers in the same state also finds that more experienced teachers tend to reduce student absences. In addition, the paper finds evidence of more experienced ELA teachers inducing more free time reading in their students, and more experienced math teachers causing students to be less likely to be victims of disciplinary offenses. Finally, more recent⁴³ studies replicate this early-career improvement, as measured by rubric-based classroom observation scores.⁴⁴

Prior work investigating what seems to influence teacher improvement over time has found evidence of teachers learning from each other,⁴⁵ learning faster when consistently assigned to the same grade,⁴⁶ and learning faster in more supportive professional environments.⁴⁷ Recent evidence⁴⁸ suggests that teachers improve in classroom management over time and that teachers who struggle with managing their classroom are more likely to exit the profession early on. In addition, teachers who remain in the teacher workforce tend to see early-career improvements in classroom management.

Regarding other readily observable teacher characteristics, a set of studies examines the relationship between teaching effectiveness and holding a graduate degree. These studies almost universally find no relationship.⁴⁹ One exception is a study from North Carolina that finds that middle- and high-school mathematics teachers with an “in-area” graduate degree are more effective than teachers with an undergraduate degree only.⁵⁰ In addition, several studies examine the impact of completing the National Board Certification (NBC) process. One study uses data from elementary-school teachers in North Carolina and finds that the NBC process identifies more effective teachers; however, the act of being certified does not increase effectiveness.⁵¹ Another



study finds similar results in Washington⁵² and another in Florida, although the latter finds less consistent evidence of National Board Certified teachers being more effective than non-National Board Certified teachers.⁵³ Finally, a paper from North Carolina finds that high-school teachers become more effective after they become certified.⁵⁴

Endnotes and references

- ¹ NCES Fast Facts. <https://nces.ed.gov/fastfacts/display.asp?id=28>.
- ² A total of 3.8 million public school teachers <https://nces.ed.gov/programs/coe/indicator/clr/public-school-teachers>) and a U.S. labor force of 168 million workers (<https://www.bls.gov/news.release/empsit.t01.htm>).
- ³ K-12 spending represents approximately 3.25% of GDP. Katz, N., K. W. Apfelbaum, S. Frank, and K. H. Miles. 2018. Low Teacher Salaries 101. Education Resource Strategies 2-16. [https://www.erstrategies.org/wp-content/uploads/2023/12/Low Teacher Salaries 101 - Updated 11.30.18.pdf](https://www.erstrategies.org/wp-content/uploads/2023/12/Low_Teacher_Salaries_101_-_Updated_11.30.18.pdf). Expenditures on staff salaries represent 55% of expenditures, two-thirds of which support instruction. NCES Fast Facts: Expenditures. <https://nces.ed.gov/fastfacts/display.asp?id=66#:~:text=Altogether%2C%20salaries%20and%20benefits%20combined,on%20tuition%20and%20other%20expenditures>.
- ⁴ Goldhaber, Dan. 2016. In Schools, Teacher Quality Matters Most: Today's Research Reinforces Coleman's Findings. Education Next 16(2): 56-63. <https://www.educationnext.org/in-schools-teacher-quality-matters-most-coleman/>.
- ⁵ Sawchuk, Stephen. 2017. Is Teacher Recertification Broken? Education Week December 6. <https://www.edweek.org/teaching-learning/is-teacher-recertification-broken/2017/12>.
- ⁶ Cowan, J., D. Goldhaber, and R. Theobald. 2020. Performance Evaluations as a Measure of Teacher Effectiveness When Standards Differ: Accounting for Variation across Classrooms, Schools, and Districts. National Center for the Analysis of Longitudinal Data in Education Research. <https://caldercenter.org/sites/default/files/CALDER%20WP%20197-0618-2.pdf>.
- ⁷ Bacher-Hicks, Andrew, and Cory Koedel. 2023. Estimation and Interpretation of Teacher Value Added in Research Applications. Handbook of the Economics of Education 6: 93-134. <https://www.sciencedirect.com/science/article/abs/pii/S1574069222000022>.
- ⁸ See, for example, the roundtable discussion by Raj Chetty, Jesse Rothstein, and Eric Hanushek on the use of VA in policy in <https://hanushek.stanford.edu/sites/default/files/publications/Hanushek%202013%20Focus%2029%282%29.pdf>.
- ⁹ Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2014. Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates. American Economic Review 104(9): 2593-2632. <https://rajchetty.com/wp-content/uploads/2021/04/w19423.pdf>.
- ¹⁰ Bacher-Hicks, Andrew, Thomas J. Kane, and Douglas O. Staiger. 2014. Validating Teacher Effect Estimates Using Changes in Teacher Assignments in Los Angeles (No. w20657). National Bureau of Economic Research. https://bpb-us-e1.wpmucdn.com/sites.dartmouth.edu/dist/9/2108/files/2019/06/Bahcer-Hicks_Kane_Staiger_Validating_Teacher_Effects_2016_text_figures_final.pdf.
- ¹¹ Rothstein, Jesse. 2017. Measuring the Impacts of Teachers: Comment. American Economic Review 107(6): 1656-1684. <https://irle.berkeley.edu/wp-content/uploads/2017/01/Revisiting-the-Impacts-of-Teachers.pdf>.
- ¹² Bacher-Hicks, Andrew, Mark J. Chin, Thomas J. Kane, and Douglas O. Staiger. 2019. An Experimental Evaluation of Three Teacher Quality Measures: Value-Added, Classroom Observations, and Student Surveys. Economics of Education Review 73: 101919. <https://sci-hub.se/https://doi.org/10.1016/j.econedurev.2019.101919>; Kane, Thomas J., and Douglas O. Staiger. 2008. Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation (No. w14607). National Bureau of Economic Research. https://www.nber.org/system/files/working_papers/w14607/w14607.pdf; Kane, Thomas J., Daniel F. McCaffrey, Trey Miller, and Douglas O. Staiger. 2013. Have We Identified Effective Teachers? Validating Measures of Effective Teaching Using Random Assignment (Research Paper). MET Project. Bill & Melinda Gates Foundation. <https://files.eric.ed.gov/fulltext/ED540959.pdf>.
- ¹³ For a full review, see Redding, Christopher. 2019. A Teacher Like Me: A Review of the Effect of Student-Teacher Racial/Ethnic Matching on Teacher Perceptions of Students and Student Academic and Behavioral Outcomes. Review of Educational Research 89(4): 499-535. <https://journals.sagepub.com/doi/abs/10.3102/0034654319853545>.



- ¹⁴ Gershenson, Seth, Cassandra M. D. Hart, Joshua Hyman, Constance A. Lindsay, and Nicholas W. Papageorge. 2022. The Long-Run Impacts of Same-Race Teachers. *American Economic Journal: Economic Policy* 14(4): 300–342. https://www.nber.org/system/files/working_papers/w25254/w25254.pdf.
- ¹⁵ Goldhaber, Dan, Vanessa Quince, and Roddy Theobald. 2018. Has It Always Been This Way? Tracing the Evolution of Teacher Quality Gaps in US Public Schools. *American Educational Research Journal* 55(1): 171–201. <https://files.eric.ed.gov/fulltext/EJ1167060.pdf>.
- ¹⁶ Isenberg, Eric, Jeffrey Max, Philip Gleason, and Jonah Deutsch. 2022. Estimating the "Effective Teaching Gap": Students Experience Unequal Outcomes, but Mostly Equal Access to High-Quality Instruction. *Education Next* 22(4): 60–65. https://www.educationnext.org/wp-content/uploads/2022/12/ednext_XXII_4_isenberg_et_al.pdf.
- ¹⁷ Jacob, Brian A., and Lars Lefgren. 2008. Can Principals Identify Effective Teachers? Evidence on Subjective Performance Evaluation in Education. *Journal of Labor Economics* 26(1): 101–136.
- ¹⁸ Backes, Ben, James Cowan, Dan Goldhaber, and Roddy Theobald. 2022. *Teachers and School Climate: Effects on Student Outcomes and Academic Disparities* (No. 274-1022). CALDER Working Paper.
- ¹⁹ Jackson, C. Kirabo. 2018. What Do Test Scores Miss? The Importance of Teacher Effects on Non-Test Score Outcomes. *Journal of Political Economy* 126(5): 2072–2107. https://bpb-us-e1.wpmucdn.com/sites.northwestern.edu/dist/b/3664/files/2020/10/FULL_paper_FINAL_stamped.pdf.
- ²⁰ Gershenson, Seth. 2016. Linking Teacher Quality, Student Attendance, and Student Achievement. *Education Finance and Policy* 11(2): 125–149. <https://direct.mit.edu/edfp/article/11/2/125/10241/Linking-Teacher-Quality-Student-Attendance-and>.
- ²¹ Liu, Jing, and Susanna Loeb. 2021. Engaging Teachers: Measuring the Impact of Teachers on Student Attendance in Secondary School. *Journal of Human Resources* 56(2): 343–379. <https://jhr.uwpress.org/content/early/2019/07/02/jhr.56.2.1216-8430R3>.
- ²² Liu, Jing, Michael Hayes, and Seth Gershenson. 2021. From Referrals to Suspensions: New Evidence on Racial Disparities in Exclusionary Discipline. <https://repec.iza.org/dp14619.pdf>.
- ²³ Backes, Ben, James Cowan, Dan Goldhaber, and Roddy Theobald. 2023. *How to Measure a Teacher: The Influence of Test and Nontest Value-Added on Long-Run Student Outcomes* (Working Paper No. 270-0423-2). National Center for Analysis of Longitudinal Data in Education Research (CALDER). <https://caldercenter.org/sites/default/files/CALDER%20WP%20270-0423-2.pdf>.
- ²⁴ Goldhaber, Dan, James Cowan, and Joe Walch. 2013. Is a Good Elementary Teacher Always Good? Assessing Teacher Performance Estimates across Subjects. *Economics of Education Review* 36: 216–228.
- ²⁵ Kraft, Matthew A. 2019. Teacher Effects on Complex Cognitive Skills and Social-Emotional Competencies. *Journal of Human Resources* 54(1): 1–36.
- ²⁶ See, for example, here and the citations within Rose, Evan K., Jonathan T. Schellenberg, and Yotam Shem-Tov. 2022. *The Effects of Teacher Quality on Adult Criminal Justice Contact* (No. w30274). National Bureau of Economic Research. https://www.nber.org/system/files/working_papers/w30274/w30274.pdf.
- ²⁷ Chamberlain, Gary E. 2013. Predictive Effects of Teachers and Schools on Test Scores, College Attendance, and Earnings. *Proceedings of the National Academy of Sciences* 110(43): 17176–17182. <https://scholar.harvard.edu/sites/scholar.harvard.edu/files/chamberlain/files/pnas-2013-chamberlain-17176-82.pdf>.
- ²⁸ Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2014. Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood. *American Economic Review* 104(9): 2633–2679. <https://rajchetty.com/wp-content/uploads/2021/04/w19424-2.pdf>.
- ²⁹ Backes et al. (2023).



- ³¹ Rockoff, J. E., Jacob, B. A., Kane, T. J., & Staiger, D. O. (2011). Can you recognize an effective teacher when you recruit one?. *Education finance and Policy*, 6(1), 43-74.
- ³² Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2007). Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Economics of education review*, 26(6), 673-682.
- ³³ Boyd, Donald, Hamilton Lankford, Susanna Loeb, Jonah Rockoff, and James Wyckoff. 2008. The Narrowing Gap in New York City Teacher Qualifications and Its Implications for Student Achievement in High-Poverty Schools (No. w14021). National Bureau of Economic Research. <https://core.ac.uk/download/pdf/6716092.pdf>.
- ³⁴ Rockoff, Jonah E. 2004. The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data. *American Economic Review* 94(2): 247-252.
- ³⁵ Harris, Douglas N., and Tim R. Sass. 2011. Teacher Training, Teacher Quality and Student Achievement. *Journal of Public Economics* 95(7-8): 798-812. https://caldercenter.org/sites/default/files/1001059_Teacher_Training.pdf.
- ⁴⁰ Kini, Tara, and Anne Podolsky. 2016. Does Teaching Experience Increase Teacher Effectiveness? A Review of the Research. Learning Policy Institute. https://learningpolicyinstitute.org/sites/default/files/product-files/Teaching_Experience_Report_June_2016.pdf.
- ⁴¹ Gershenson (2016).
- ⁴² Ladd, Helen F., and Lucy C. Sorensen. 2017 Returns to Teacher Experience: Student Achievement and Motivation in Middle School. *Education Finance and Policy* 12(2): 241-279. https://www.researchgate.net/profile/Lucy-Sorensen/publication/301247657_Returns_to_Teacher_Experience_Student_Achievement_and_Motivation_in_Middle_School/links/59f9aa730f7e9b553ec0f81a/Returns-to-Teacher-Experience-Student-Achievement-and-Motivation-in-Middle-School.pdf.
- ⁴³ Bell, Courtney A., Jessalynn K. James, Eric S. Taylor, and James Wyckoff. 2023. Measuring Returns to Experience Using Supervisor Ratings of Observed Performance: The Case of Classroom Teachers (No. w30888). National Bureau of Economic Research. https://scholar.harvard.edu/files/erictaylor/files/returns-exper-obsv-ratings-bjtw_jan-23.pdf.
- ⁴⁴ Kraft, Matthew A., John P. Papay, and Olivia L. Chi. 2020. Teacher Skill Development: Evidence from Performance Ratings by Principals. *Journal of Policy Analysis and Management* 39(2): 315-347. https://scholar.harvard.edu/files/mkraft/files/kraft_papay_chi_2018_teacher_skill_development.pdf.
- ⁴⁵ Jackson, C. Kirabo, and Elias Bruegmann. 2009. Teaching Students and Teaching Each Other: The Importance of Peer Learning for Teachers. *American Economic Journal: Applied Economics* 1(4): 85-108. https://bpb-us-e1.wpmucdn.com/sites.northwestern.edu/dist/b/3664/files/2020/10/fulltext_stamped-1.pdf.
- ⁴⁶ Ost, Ben. 2014. How Do Teachers Improve? The Relative Importance of Specific and General Human Capital. *American Economic Journal: Applied Economics* 6(2): 127-151. <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=48e8e401e853d74421d152cfe239c1c97cb24583>.
- ⁴⁷ Kraft, Matthew A., and John P. Papay. 2014. Can Professional Environments in Schools Promote Teacher Development? Explaining Heterogeneity in Returns to Teaching Experience. *Educational Evaluation and Policy Analysis* 36(4): 476-500. https://scholar.harvard.edu/sites/scholar.harvard.edu/files/mkraft/files/kraft_papay_-_prof_env_teacher_development_eepa_full.pdf.
- ⁴⁸ Bartanen, Brendan, Courtney Bell, Jessalynn James, Eric S. Taylor, and James H. Wyckoff. "Refining" Our Understanding of Early Career Teacher Skill Development: Evidence From Classroom Observations. <https://edworkingpapers.com/sites/default/files/ai23-845.pdf>.



⁴⁹ For example, Rivkin et al. (2005) noted above and Chingos, Matthew M., and Paul E. Peterson. 2011. It's Easier to Pick a Good Teacher than to Train One: Familiar and New Results on the Correlates of Teacher Effectiveness. *Economics of Education Review* 30(3): 449–465. <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=8c63564c06becf5e047b38e0ed08c8567e96c10b8ref=pmp-magazine.com>.

⁵⁰ Bastian, Kevin C. 2019. A Degree Above? The Value-Added Estimates and Evaluation Ratings of Teachers with a Graduate Degree. *Education Finance and Policy* 14(4): 652–678. <https://direct.mit.edu/edfp/article/14/4/652/12331>.

⁵¹ Goldhaber, Dan, and Emily Anthony. 2007. Can Teacher Quality Be Effectively Assessed? National Board Certification as a Signal of Effective Teaching. *The Review of Economics and Statistics* 89(1): 134–150. <https://www.jstor.org/stable/40043080>.

⁵² Cowan, James, and Dan Goldhaber. 2016. National Board Certification and Teacher Effectiveness: Evidence from Washington State. *Journal of Research on Educational Effectiveness* 9(3): 233–258. <https://www.tandfonline.com/doi/pdf/10.1080/19345747.2015.1099768>.

⁵³ Harris, Douglas N., and Tim R. Sass. 2009. The Effects of NBPTS-Certified Teachers on Student Achievement. *Journal of Policy Analysis and Management: The Journal of the Association for Public Policy Analysis and Management* 28(1): 55–80. <https://www.jstor.org/stable/pdf/29738986.pdf>.

⁵⁴ Clotfelter, Charles T., Helen F. Ladd, and Jacob L. Vigdor. 2010. Teacher Credentials and Student Achievement in High School: A Cross-Subject Analysis with Student Fixed Effects. *Journal of Human Resources* 45(3): 655–681. <https://jhr.uwpress.org/content/wpjhr/45/3/655.full.pdf>.

Suggested citation

Backes, Ben (2025). "The Importance of Teachers," in *Live Handbook of Research on Education Finance and Policy*, in Douglas Harris (ed.), Association for Education Finance and Policy, viewed MM/DD/YYYY, <WEB LINK>.



live handbook



AEPF

The Importance of Teachers | Ben Backes